

# Simple Closed-Form Approximations for Achievable Information Rates of Coded Modulation Systems

Maria Urlea and Sergey Loyka 

**Abstract**—The intuitive sphere-packing argument is used to obtain analytically-tractable closed-form approximations for achievable information rates of coded modulation transmission systems, for which only analytically-intractable expressions are available in the literature. These approximations provide a number of insights, possess useful properties, and facilitate design/optimization of such systems. They apply to constellations of various cardinalities (including large ones), are simple yet reasonably accurate over the whole signal-to-noise ratio range, and compare favorably to the achieved rates of recent state-of-the-art experiments.

**Index Terms**—Achievable information rate (AIR), quadrature amplitude modulation (QAM), spectral efficiency, approximation.

## I. INTRODUCTION

THE growing demand for higher transmission rates and higher spectral efficiency along with the development of powerful capacity-approaching codes has recently stimulated significant interest in coded-modulation optical transmission systems, with multi-level modulation and capacity-approaching codes [1], [2]. For such systems, an adequate performance metric is achievable information rate (AIR), which can be expressed as mutual information (MI) or generalized mutual information (GMI), depending on the specifics of decoding/demodulation algorithms [3], [4]. They have been widely used for analysis, optimization and design of various communication systems, including optical fiber systems.

While achievable information rates have been studied in the information-theoretic literature for multi-level modulation formats over the additive white Gaussian noise (AWGN) channel [22]–[24], the corresponding expressions cannot be evaluated analytically in a closed form, even for the simplest AWGN channel, since they include a number of analytically-intractable integrals over infinite intervals. Numerical integration (e.g., via Monte-Carlo approach) remains the only alternative [4], [22]. This makes it difficult to obtain insights and also to perform system design and optimization (for example, to find an optimal power allocation in a multi-stream transmission system, as in e.g. [25]).

Manuscript received August 11, 2020; revised October 16, 2020; accepted November 5, 2020. Date of publication November 18, 2020; date of current version March 1, 2021. (Corresponding author: Sergey Loyka.)

The authors are with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada (e-mail: murle095@uottawa.ca; sergey.loyka@uottawa.ca).

Color versions of one or more of the figures in this article are available at <https://doi.org/10.1109/JLT.2020.3039178>.

Digital Object Identifier 10.1109/JLT.2020.3039178

In this paper, we present analytically-tractable closed-form approximations to achievable information rates for  $M$ -PAM and  $M$ -QAM constellations (modulation) on the AWGN channel (possibly including nonlinear interference in optical fiber modeled as additional Gaussian noise [5]) using the intuitive sphere-packing argument. These approximations are simple yet reasonably accurate over the whole signal-to-noise ratio (SNR) range (not just at high or low SNR) and for various constellation cardinalities  $M$  (including large ones). Comparison to the back-to-back (B2B) or long-haul end-to-end (E2E) rates of recent state-of-the-art systems/experiments shows that the proposed approximations are in fact more accurate than the ideal AIR (via MI or GMI) since the latter ignores many imperfections and limitations of real-world systems (e.g. non-zero overhead of realistic codes, guard bands, etc.).

The proposed approximations also provide a number of insights unavailable from the exact (MI-based) expressions. In particular, they allow one to obtain the minimum  $M$  required to approach closely the (modulation-unconstrained) AWGN channel capacity without using unnecessarily large constellations. The derivation of these approximations offers an additional insight into a mechanism causing a rate loss in the modulation-constrained system: while the rate loss is additive at low SNR, it is multiplicative, i.e., much more pronounced, at high SNR. The obtained approximations possess useful properties: they are differentiable and concave in the SNR so that the respective optimization problems are convex and thus all powerful tools of convex optimization (see e.g. [28]) can be used to solve them. The main contributions of this paper are the approximations of the modulation-constrained achievable information rates of  $M$ -PAM and  $M$ -QAM constellations in (11), (17), and (19), from which the minimum number of constellation points to approach closely the channel capacity (without significant rate loss due to limited  $M$ ) can be found as in (12) and (18).

## II. CHANNEL MODEL AND INFORMATION RATES

Following the widely-accepted approach, we use the Gaussian channel model, which has been extensively used in many areas of digital communications [26]. This model was also demonstrated (via simulations and also experimentally) to be sufficiently accurate for the optical fiber channel in many cases of practical interest (uncompensated transmission in the low-to-moderate nonlinearity regime) and offers a good balance of accuracy and simplicity [5], [6]. The system block diagram is shown in Fig. 1,

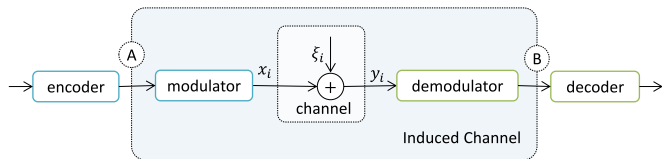


Fig. 1. System model and induced channel.

which includes an encoder, a modulator, a channel, a demodulator and a decoder [26]. The function of the encoder is to protect the transmitted message against channel-induced errors using forward error correction and the function of the modulator is to transform the encoded signal to a form suitable for transmission over the communication channel. The baseband discrete-time channel model (after matched filtering and sampling) is

$$y_i = x_i + \xi_i, \quad (1)$$

where  $x_i$  and  $y_i$  are transmitted and received (real-valued) symbols respectively at time  $i$ ,  $\xi_i$  is the additive white Gaussian noise. The capacity of this channel, i.e. the maximum achievable rate under the reliability criterion (arbitrary small error probability) and the power constraint, is [23]

$$C = \frac{1}{2} \log(1 + \gamma) \text{ [bit/symbol]}, \quad (2)$$

where  $\gamma = \sigma_x^2/\sigma_0^2$  is the SNR,  $\sigma_x^2$  and  $\sigma_0^2$  are the signal and noise power and all logarithms are base-2;  $\sigma_0^2$  can also include nonlinear interference in optical fiber when modeled as Gaussian noise, see e.g. [5].

In practical systems, digital modulator makes use of a certain number of signals conveniently represented via a constellation (in the signal space) [26]. Due to implementation complexity, the constellation cardinality (i.e. the number of points/signals) is limited to  $M$ . On the other hand, recent development and adoption of powerful capacity-approaching codes allows one to build nearly-optimal encoders and decoders. Consequently, modulation-constrained achievable information rates are widely used as a performance metric [1]–[4]. From the information-theoretic perspective, the part of the system between the encoder and the decoder can be interpreted as an *induced channel*, which includes the actual channel and the constrained (fixed) modulator/demodulator. This is illustrated in Fig. 1, where the induced channel is between A and B. Assuming symbol-wise decoding, the achievable information rate of this system is the mutual information of the induced channel. This quantity has been studied in the information-theoretic literature [22]–[24]. For the real-valued  $M$ -PAM constellation under equiprobable signalling, the per-symbol MI  $C_M$  is [23]

$$C_M = \log M$$

$$- \sum_j \frac{1}{M \sqrt{2\pi\sigma_0^2}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2\sigma_0^2}} \log \sum_i e^{-\frac{d_{ij}^2}{2\sigma_0^2}} e^{-\frac{z d_{ij}}{\sigma_0^2}} dz, \quad (3)$$

where  $d_{ij} = a_i - a_j$  so that  $|d_{ij}|$  is the distance between constellation points  $a_i$  and  $a_j$ ; the summation is over all points. Unfortunately, the integrals over infinite intervals cannot be evaluated in closed-form, which makes it difficult to obtain insights and to use it in design/optimization process. Numerical evaluation of these integrals can be also difficult, especially when constellation cardinality is large and/or real time evaluation is required (e.g. to compute optimal power allocation for multi-stream transmission system, see e.g. [25]).

From the information-theoretic perspective, achieving the channel capacity requires selecting the best encoder/decoder and modulator/demodulator while a modulation-constrained achievable information rate implies the optimization of the encoder/decoder only while the modulator is fixed, so that  $C_M \leq C$ . On the other hand, the AIR cannot exceed the entropy of the corresponding constellation [23],  $C_M \leq \log M$ , so that  $C_M$  can be upper-bounded as follows

$$C_M \leq \min(C, \log M). \quad (4)$$

While the former two bounds can be rather loose individually (at high and low SNR respectively), the latter is significantly tighter when considered over the whole SNR range (see Fig. 4). To obtain further insights and to overcome the above-mentioned difficulties, we present below a simple closed-form approximation for  $C_M$ .

### III. AIR VIA THE SPHERE PACKING ARGUMENT

While a rigorous information-theoretic derivation of (2) is available and well-known [23], the sphere-packing argument, which originates back to Shannon, provides a more intuitive understanding as to why this expression holds [27]. Since our approximation is also derived via this argument, we briefly review it below; see [27] for more details.

#### A. Sphere Packing

During the transmission,  $n$  consecutive symbols are grouped into codewords, so that  $\mathbf{x} = [x_1, x_2, \dots, x_n]$  is transmitted while  $\mathbf{y} = \mathbf{x} + \boldsymbol{\xi}$  is received, where  $\boldsymbol{\xi} = [\xi_1, \xi_2, \dots, \xi_n]$  is the noise vector. As  $n$  is getting large, the total noise power  $|\boldsymbol{\xi}|^2 = \sum_i \xi_i^2$  approaches  $n\sigma_0^2$  with high probability (due to the law of large numbers), which is known as “sphere hardening” [27], so that  $\mathbf{y}$  belongs to the noise sphere centered on a transmitted codeword, as shown in Fig. 2, with high probability. The decoder decides in favor of a particular codeword when received signal  $\mathbf{y}$  belongs to the codeword region of that codeword (also known as “decoding region”). For any transmitted codeword, the received signal  $\mathbf{y}$  belongs with high probability to the received signal sphere of radius  $\sqrt{n(\sigma_x^2 + \sigma_0^2)}$  which encloses all individual noise spheres centered on respective codewords. For reliable decoding, (i) codeword regions must enclose corresponding noise spheres, and (ii) they must not overlap. These can be satisfied by selecting codeword regions as spheres of a radius slightly larger than that of noise spheres and by packing all those spheres into the received signal sphere as tightly as possible allowing no overlaps.

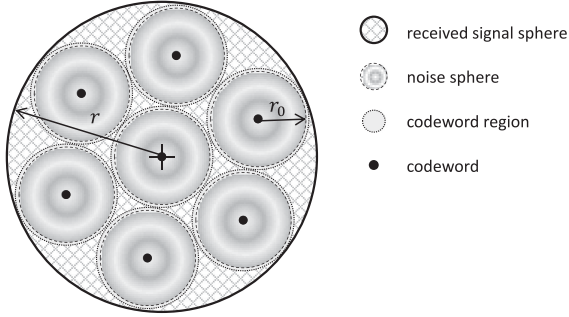


Fig. 2. Sphere-packing argument: a codeword region is just slightly larger than a corresponding noise sphere;  $r_0 \approx \sqrt{n\sigma_0^2}$ ,  $r \approx \sqrt{n(\sigma_x^2 + \sigma_0^2)}$ .

The number  $N$  of distinct codewords that can be reliably transmitted over the channel is approximately equal to the number of non-overlapping noise spheres that can fit into the received signal sphere. It can be estimated from the ratio of respective volumes:

$$N \approx V/V_0 = (1 + \gamma)^{n/2}, \quad (5)$$

where  $V = \alpha(n\sigma_x^2 + n\sigma_0^2)^{n/2}$  and  $V_0 = \alpha(n\sigma_0^2)^{n/2}$  are the volumes of the received signal and noise spheres respectively,  $\alpha = \pi^{n/2}/\Gamma(n/2 + 1)$  and  $\Gamma(\cdot)$  is the gamma function, so that the channel capacity is

$$C \approx \frac{1}{n} \log N \approx \frac{1}{2} \log(1 + \gamma). \quad (6)$$

Remarkably, this heuristic argument gives the exact value of the channel capacity for the AWGN channel.

### B. Achievable Information Rates

The argument above applies if the number of codewords of length  $n$  can be as large as necessary (i.e. is not constrained). For a fixed constellation of  $M$  points, the number of codewords of length  $n$  can be at most  $M^n$ . To evaluate the impact of this constraint, we present the per-symbol AIR in the following form:

$$C_M = C - \Delta C, \quad (7)$$

where  $C$  is the (unconstrained) channel capacity (as above) and  $\Delta C \geq 0$  is the rate loss due to a fixed constellation of cardinality  $M$ . To estimate the latter, consider a hypothetical system with noise power  $\sigma_1^2$  such that the number of distinct codewords (noise spheres) is exactly  $M^n$ :

$$M^n = \frac{V}{V_1} = \frac{(n\sigma_x^2 + n\sigma_1^2)^{n/2}}{(n\sigma_1^2)^{n/2}}, \quad (8)$$

where  $V_1 = \alpha(n\sigma_1^2)^{n/2}$  is the volume of the hypothetical noise sphere, which is also the volume of a codeword region when there are exactly  $M^n$  codewords. For this system, there is no loss in capacity due to a fixed constellation (within the sphere packing approximation) since the noise power is “right” (i.e. noise spheres are the same as respective codeword regions so that no more codewords can fit without increasing the error

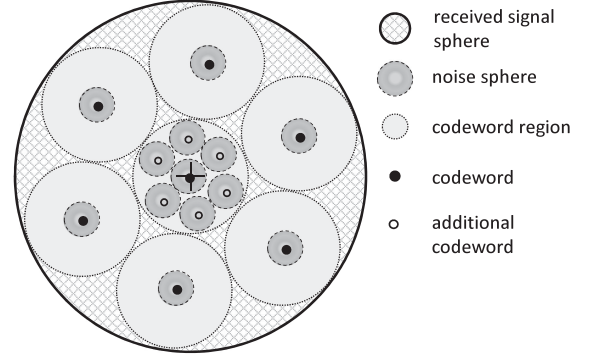


Fig. 3. The impact of the limited number of codewords: more noise spheres representing additional codewords could be packed into existing codeword regions when noise is small.

probability):

$$\sigma_1^2 = \frac{\sigma_x^2}{M^2 - 1} \approx \frac{\sigma_x^2}{M^2}, \quad (9)$$

where the last approximation holds when  $M$  is reasonably large,  $M^2 \gg 1$ . However, if the true noise power is less than the hypothetical one,  $\sigma_0^2 < \sigma_1^2$ , more than one noise sphere can fit within the hypothetical noise sphere (codeword region), as shown in Fig. 3 (the central sphere). The resulting  $\Delta C$  can be interpreted as the capacity of the fictitious channel with signal power  $\sigma_1^2$  and the noise power  $\sigma_0^2$ , which can be estimated via the ratio of volumes approach similarly to (6):

$$\Delta C \approx \frac{1}{n} \log \frac{(n\sigma_1^2 + n\sigma_0^2)^{n/2}}{(n\sigma_0^2)^{n/2}} \approx \frac{1}{2} \log \left( 1 + \frac{\sigma_x^2}{M^2\sigma_0^2} \right). \quad (10)$$

Substituting this in (7), one finally obtains an approximation  $C_a$  for the per-symbol AIR of  $M$ -PAM:

$$C_M \approx C_a = \frac{1}{2} \log \frac{1 + \gamma}{1 + \gamma/M^2}. \quad (11)$$

While this approximation is simple and analytically tractable, it is also reasonably accurate over the whole SNR range and for various constellation cardinalities  $M$  (especially when  $M$  is large), as Fig. 4 shows, being somewhat less accurate in the transition region between low and high SNR regimes and for 2-PAM (for which more accurate approximation is obtained in the next section). In all considered cases, the approximation inaccuracy does not exceed about 0.3 b/symb, which transforms into 30% for 2-PAM, 10% and 5% for 8-PAM and 64-PAM, respectively. Hence, the approximation is more accurate for higher-rate constellations and also significantly more accurate outside of the transition region (between low and high SNR). Since the experimental results are well below the ideal AIR, this approximation is more accurate than the ideal AIR, with respect to the rates of real-world systems/experiments (see also the next section).

Not only this approximation is much simpler than its exact counterpart in (3) and thus facilitates a numerical or even analytical optimization (e.g. to find an optimal power allocation in a multi-stream transmission system, for which no closed-form

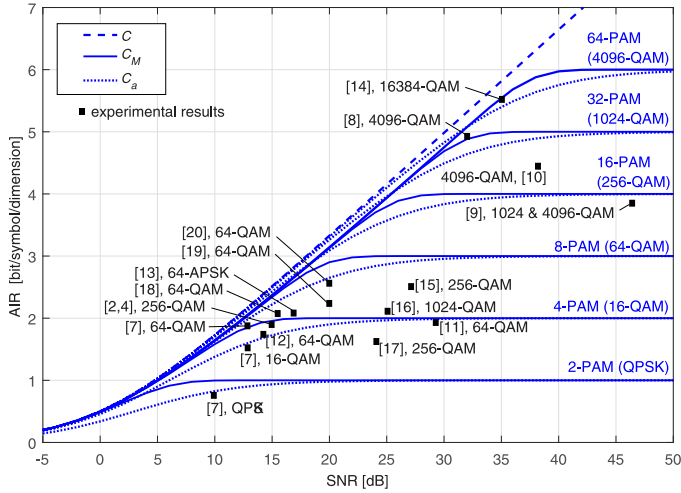


Fig. 4.  $M$ -PAM (QAM) AIR in [b/sym./dimension] evaluated via numerical integration in (3), its approximation in (11) (or (17)) and the AWGN channel capacity in (2) vs. SNR; recent experimental results are also shown for comparison.

solution is known [25]), but it also captures important properties of the ideal AIR  $C_M$  and provides additional insights unavailable from the known results:

1) As  $M$  increases, both  $C_a$  and  $C_M$  approach  $C$ :  $C_a \approx C_M \approx C$  if  $M \gg \sqrt{\gamma}$ , from which one can estimate minimum  $M$  required to approach closely the channel capacity without using unnecessarily large constellations:<sup>1</sup>

$$M_{\min} \approx 2 \max\{1, \sqrt{\gamma}\}. \quad (12)$$

In this regime, the upper bound in (4) is tight.

2) At high SNR, both  $C_a$  and the ideal AIR  $C_M$  approach the constellation entropy and are independent of the SNR,

$$C_a \approx C_M \approx \log M \text{ if } \gamma \gg M^2, \quad (13)$$

so that the upper bound in (4) is also tight at high SNR.

3) At low SNR,  $C_a$  can be approximated as follow:

$$C_a \approx \left(1 - \frac{1}{M^2}\right) \frac{\log e}{2} \gamma \text{ if } \gamma \ll 1, \quad (14)$$

so that the rate loss is small if  $M^2 \gg 1$  and never exceeds 25%. Comparing this to (13), one concludes that the rate loss (due to constrained modulation) is additive at low SNR and multiplicative at high SNR, i.e. much more pronounced in the latter case.

4) The approximation  $C_a$  is a continuously-differentiable (to any order) and increasing function of the SNR:

$$\frac{\partial C_a}{\partial \gamma} = \frac{\log e}{2} \frac{M^2 - 1}{(1 + \gamma)(M^2 + \gamma)} > 0. \quad (15)$$

<sup>1</sup>This approximation demonstrates that the upper bound in [24], which can be put in the form  $M_{\min} \leq 2\sqrt{1 + \gamma}$ , is actually tight. Note also that the upper bound was obtained in [24] via an elaborate information-theoretic analysis (which does not yield a capacity approximation) while our approximation to  $M_{\min}$  follows directly from the approximation in (11).

5)  $C_a$  is a strictly-concave function of the SNR:

$$\frac{\partial^2 C_a}{\partial \gamma^2} = \frac{\log e}{2} \frac{(1 - M^2)(1 + M^2 + 2\gamma)}{(1 + \gamma)^2(M^2 + \gamma)^2} < 0. \quad (16)$$

The latter two properties are instrumental in numerical optimization of multi-stream transmission systems (e.g. MIMO, as in [25]) by rendering the respective optimization problems convex so that all powerful tools of convex optimization, see e.g. [28], can be used.

### C. $M$ -QAM

While the approximation above was derived for the  $M$ -PAM constellation (which is 1-D), it can also be extended to  $M$ -QAM (2-D), which is a popular choice for modern optical transmission systems [1], [2] as well as wireless systems (e.g. WiFi, LTE, 5G). Indeed,  $M$ -QAM can be considered as  $2 \times \sqrt{M}$ -PAM operating on in-phase and quadrature channels [26] (here we assume that  $\sqrt{M}$  is integer), so that its per-symbol AIR can be approximated as follows:

$$C_{M-QAM} = 2C_{\sqrt{M}-PAM} \approx \log \frac{1 + \gamma}{1 + \gamma/M}, \quad (17)$$

and the properties above hold with the substitution  $M \rightarrow \sqrt{M}$ . In particular, the minimum constellation cardinality to approach closely the channel capacity is

$$M_{\min} \approx 4 \max\{1, \gamma\} \quad (18)$$

Since 1st equality in (17) holds for both the approximations and ideal AIR, the relative accuracy of the  $M$ -QAM approximation is the same as that of  $M$ -PAM, see Fig. 4, where the rates are in [bit/symbol/dimension] (recall that  $M$ -PAM is 1-D while  $M$ -QAM is 2-D; experimental results are 4-D since two polarizations are used). Comparing this approximation with recent state-of-the-art experimental results<sup>2</sup> for back-to-back (B2B) or long-haul end-to-end (E2E) rates in Fig. 4, one concludes that the approximation is in fact closer to the experimental rates than the ideal AIR. Note a particularly good agreement for QPSK, 16-QAM and 64-QAM in [7], and 64-QAM in [20]. We attribute this to the fact that the approximation  $C_{M-QAM}$  in (17) underestimates the ideal AIR  $C_M$  in (3) while the experimental B2B or E2E rates are also below the ideal AIR by about 0.5-2 [b/sym.] (due to finite blocklength/complexity codes with non-zero overhead and non-zero guard bands, in addition to other imperfections of practical systems [21]) Hence, the proposed approximations are more accurate than the ideal AIR  $C_M$  with respect to the rates of real-world systems.

## IV. AN ASYMPTOTIC APPROXIMATION

In order to improve the accuracy of the AIR approximations above for BPSK and QPSK, we obtain here an approximation via the tools of asymptotic analysis (see e.g. [29] for more details on these tools). Applying the Laplace method to the integrals

<sup>2</sup>It should be noted that the quoted experimental results were obtained under different conditions, e.g. single-channel vs. multi-channel systems etc., and hence differ significantly.

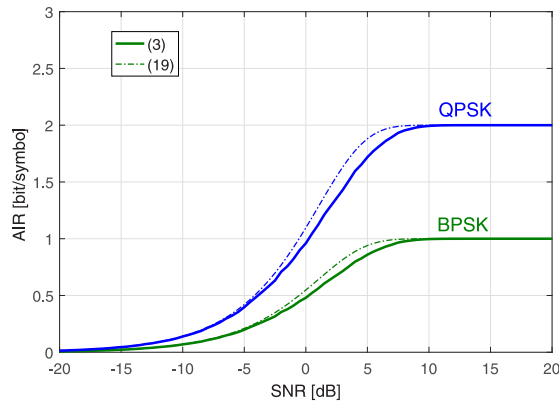


Fig. 5. 2-PAM (BPSK) and 4-QAM (QPSK) AIR and the approximations in (19) vs. SNR.

in (3), the AIR of 2-PAM (BPSK) and 4-QAM (QPSK) can be approximated, after some manipulations, as

$$C_{2-PAM} \approx 1 - \log(1 + e^{-\gamma}), \quad C_{4-QAM} = 2C_{2-PAM} \quad (19)$$

As Fig. 5 shows, this approximation is remarkably accurate over the whole SNR range: the inaccuracy does not exceed 10%, being much smaller outside of the transition region. Note also that this approximation is a continuously-differentiable (to any order) and concave function of the SNR, making it suitable for convex optimization algorithms.

An alternative approximation for the 2-PAM (BPSK) and 4-QAM (QPSK) rates, which is somewhat more involved due to its use of the error function, can be found in [30, (45), (46)].

## V. CONCLUSION

Based on the intuitive sphere-packing approach, the simple and analytically-tractable approximations have been obtained for the achievable information rates of  $M$ -PAM and  $M$ -QAM constellations over the AWGN channel, including optical fiber channels when nonlinearity can be modeled by additive Gaussian noise. Not only these approximations are reasonably accurate over the whole SNR range and for various values of  $M$  (including large ones), they also provide additional insights unavailable from the known expressions. The sphere-packing based derivation sheds additional light on the rate loss mechanism due to a fixed (constrained) constellation cardinality: while the loss is additive at low SNR, it is multiplicative, i.e. much more pronounced, at high SNR.

Comparison to the B2B or E2E rates of state-of-the-art real systems/experiments shows that the proposed simple approximations are in fact more accurate than the ideal AIR.

The presented approach and approximations can also be used for polarization-multiplexed transmission systems. Since the difference between the MI and GMI is small for Gray-labeled constellations [4], the above approximations can also be used for the latter.

## REFERENCES

- [1] M. Karlsson and E. Agrell, "Multidimensional modulation and coding in optical transport," *IEEE J. Lightw. Technol.*, vol. 35, no. 5, pp. 876–884, Feb. 2017.
- [2] J. Cho and P. J. Winzer, "Probabilistic constellation shaping for optical fiber communications," *IEEE J. Lightw. Technol.*, vol. 37, no. 6, pp. 1590–1607, Mar. 2019.
- [3] T. Fehenberger, A. Alvarado, P. Bayvel, and N. Hanik, "On achievable rates for long-haul fiber-optic communications," *Opt. Express.*, vol. 23, no. 7, pp. 9183–9191, Apr. 2015.
- [4] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems, "Achievable information rates for fiber optics: Applications and computations," *IEEE J. Lightw. Technol.*, vol. 36, no. 2, pp. 424–439, Jan. 2018.
- [5] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *IEEE J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb. 2014.
- [6] T. A. Eriksson, T. Fehenberger, P. A. Andrekson, M. Karlsson, N. Hanik, and E. Agrell, "Impact of 4D channel distribution on the achievable rates in coherent optical communication experiments," *IEEE J. Lightw. Technol.*, vol. 34, no. 9, pp. 2256–2266, May 2016.
- [7] J. Cho *et al.*, "Trans-atlantic field trial using high spectral efficiency probabilistically shaped 64-QAM and single-carrier real-time 250-Gb/s 16-QAM," *IEEE J. Lightw. Technol.*, vol. 36, no. 1, pp. 103–113, Jan. 2018.
- [8] S. L. Olsson, J. Cho, S. Chandrasekhar, X. Chen, P. J. Winzer, and S. Makovejs, "Probabilistically shaped PDM 4096-QAM transmission over up to 200 km of fiber using standard intradyne detection," *Opt. Express.*, vol. 26, no. 4, pp. 4522–4530, Feb. 2018.
- [9] S. Okamoto, M. Terayama, M. Yoshida, K. Kasai, T. Hirooka, and M. Nakazawa, "Experimental and numerical comparison of probabilistically shaped 4096 QAM and a uniformly shaped 1024 QAM in all-Raman amplified 160 km transmission," *Opt. Express.* vol. 26, no. 3, pp. 3535–3543, Feb. 2018.
- [10] S. L. Olsson, J. Cho, S. Chandrasekhar, X. Chen, E. C. Burrows, and P. J. Winzer, "Record-high 17.3-bit/s/Hz spectral efficiency transmission over 50 km using probabilistically shaped PDM 4096-QAM," in *Proc. Opt. Fiber Commun. Conf. Expo.*, Mar. 2018, pp. 1–3.
- [11] T. Zami *et al.*, "31.2-Tb/s real time bidirectional transmission of 78x400 Gb/s interleaved channels over C band of one 90-km SMF span," in *Proc. Opt. Fiber Commun. Conf., Opt. Soc. Amer.*, Mar. 2018, pp. 1–3.
- [12] F. Jardel *et al.*, "Exploring and experimenting with shaping designs for next-generation optical communications," *IEEE J. Lightw. Technol.*, vol. 36, no. 22, pp. 5298–5308, Sep. 2018.
- [13] S. Zhang *et al.*, "Capacity-approaching transmission over 6375 km at spectral efficiency of 8.3 bit/s/Hz," in *Proc. Opt. Fiber Commun. Conf., Opt. Soc. Amer.*, 2016, pp. 1–3.
- [14] X. Chen, J. Cho, A. Adamiecki, and P. Winzer, "16384-QAM transmission at 10 GBd over 25-km SSMF using polarization-multiplexed probabilistic constellation shaping," in *Proc. 45th Eur. Conf. Opt. Commun.*, Dublin, Ireland, Sep. 2019, pp. 1–4.
- [15] A. Matsushita, M. Nakamura, S. Yamamoto, F. Hamaoka, and Y. Kisaka, "41-Tbps C-Band WDM transmission with 10-Bps/Hz spectral efficiency using 1-Tbps/ $\lambda$  signals," *IEEE J. Lightw. Technol.*, vol. 38, no. 11, pp. 2905–2911, Jun. 2020.
- [16] F. Klejs *et al.*, "Spectrally efficient DP-1024 QAM 640 Gb/s long haul transmission using a frequency Comb," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, Mar. 2020, pp. 1–3.
- [17] T. Sasai, A. Matsushita, M. Nakamura, S. Okamoto, F. Hamaoka, and Y. Kisaka, "Laser phase noise tolerance of uniform and probabilistically shaped QAM signals for high spectral efficiency systems," *IEEE J. Lightw. Technol.*, vol. 38, no. 2, pp. 439–446, Jan. 2020.
- [18] J. X. Cai *et al.*, "Performance comparison of probabilistically shaped QAM formats and hybrid shaped APSK formats with coded modulation," *IEEE J. Lightw. Technol.*, vol. 38, no. 12, pp. 3280–3288, Mar. 2020.
- [19] M. Arikawa, M. Sato, and K. Hayashi, "Wide range rate adaptation of QAM-based probabilistic constellation shaping using a fixed FEC with blind adaptive equalization," *Opt. Express.*, vol. 28, no. 2, pp. 1300–1315, Jan. 2020.
- [20] M. Mazur, J. Schrder, M. Karlsson, and P. A. Andrekson, "Joint super-channel digital signal processing for effective inter-channel interference cancellation," *IEEE J. Lightw. Technol.*, vol. 38, no. 20, pp. 5676–5684, Jun. 2020.
- [21] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: The previous 20 and the next 20 years," *Opt. Express.*, vol. 26, no. 18, pp. 24190–24239, Sep. 2018.

- [22] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 1, pp. 55–67, Jan. 1982.
- [23] R. E. Blahut, *Principles and Practice of Information Theory*. Reading, MA, USA: Addison-Wesley, 1987.
- [24] L. H. Ozarow and A. D. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Jul. 1990.
- [25] A. Lozano, A. M. Tulino, and S. Verdu, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, Jul. 2006.
- [26] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*. Boston, MA, USA: Kluwer Academic, 2003.
- [27] J. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York, NY, USA: Wiley, 1965.
- [28] S. Boyd and L. Vandenberghe, "Convex optimization," Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [29] F. W. J. Olver, *Asymptotics and Special Functions*. Moscow: Nauka, 1990.
- [30] M. Secondini, E. Forestieri, and G. Prati, "Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps," *IEEE J. Lightw. Technol.*, vol. 31, no. 23, pp. 3839–3852, Dec. 2013.

**Sergey Loyka** was born in Minsk, Belarus. He received the M.S. degree (with Hons.) from Minsk Radioengineering Institute, Minsk, in 1992 and the Ph.D. degree in radio engineering from the Belorussian State University of Informatics and Radioelectronics (BSUIR), Minsk, in 1995. Since 2001, he has been a Faculty Member with the School of Electrical Engineering and Computer Science, University of Ottawa, Canada. Prior to that, he was a Research Fellow with the Laboratory of Communications and Integrated Microelectronics (LACIME) of Ecole de Technologie Supérieure, Montreal, Canada, a Senior Scientist with the Electromagnetic Compatibility Laboratory of BSUIR, Belarus, an Invited Scientist with the Laboratory of Electromagnetism and Acoustic (LEMA), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. His research areas are wireless communications and networks and, in particular, MIMO systems and security aspects of such systems, in which he has published extensively. He was the recipient of a number of awards from the URSI, the IEEE, the Swiss, Belarus and former USSR governments, and the Soros Foundation.

**Maria Urlea** received the B.A.Sc. and M.A.Sc. degrees in telecommunications from the School of Electrical Engineering and Computer Science, University of Ottawa, in 2012 and 2015, respectively. She is currently a Systems Engineer in Ottawa, Canada and focuses on network design and architectures for several local and national entities.